Direct interaction between uracil-DNA glycosylase and a proliferating cell nuclear antigen homolog in the crenarchaeon Pyrobaculum aerophilum

Yang, H ; Chiang, J-H ; Fitz-Gibbon, S ; Lebel, M ; Sartori, Alessandro A ; Jiricny, J ; Slupska, M M ; Miller, J H

Abstract: Proliferating cell nuclear antigen (PCNA) acts as a sliding clamp on duplex DNA. Its homologs, present in Eukarya and Archaea, are part of protein complexes that are indispensable for DNA replication and DNA repair. In Eukarya, PCNA is known to interact with more than a dozen different proteins, including a human major nuclear uracil-DNA glycosylase (hUNG2) involved in immediate postreplicative repair. In Archaea, only three classes of PCNA-binding proteins have been reported previously: replication factor C (the PCNA clamp loader), family B DNA polymerase, and flap endonuclease. In this study, we report a direct interaction between a uracil-DNA glycosylase (Pa-UDGa) and a PCNA homolog (Pa-PCNA1), both from the hyperthermophilic crenarchaeon Pyrobaculum aerophilum (T(opt) = 100 degrees C). We demonstrate that the Pa-UDGa-Pa-PCNA1 complex is thermostable, and two hydrophobic amino acid residues on Pa-UDGa (Phe(191) and Leu(192)) are shown to be crucial for this interaction. It is interesting to note that although Pa-UDGa has homologs throughout the Archaea and bacteria, it does not share significant sequence similarity with hUNG2. Nevertheless, our results raise the possibility that Pa-UDGa may be a functional analog of hUNG2 for PCNA-dependent postreplicative removal of misincorporated uracil.

DOI: https://doi.org/10.1074/jbc.M201820200

DNA Replication Repair and Recombination:

Direct interaction between uracil-DNA glycosylase and a PCNA homolog in the crenarchaeon Pyrobaculum aerophilum

Hanjing Yang, Ju-Huei Chiang, Sorel Fitz-Gibbon, Michel Lebel, Alessandro A. Sartori, Joseph Jiricny, Malgorzata M. Slupska and Jeffrey H. Miller

J. Biol. Chem. published online April 1, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M201820200

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2002/04/01/jbc.M201820200.citation.full.html#ref-list-1

JBC Papers in Press. Published on April 1, 2002 as Manuscript M201820200

Direct Interaction between Uracil-DNA Glycosylase and a PCNA Homolog in the Crenarchaeon *Pyrobaculum aerophilum*

Hanjing Yang, Ju-Huei Chiang, Sorel Fitz-Gibbon¶, Michel Lebel[1], Alessandro A. Sartori[2], Joseph Jiricny[2], Malgorzata M. Slupska‡, and Jeffrey H. Miller*

Department of Microbiology and Molecular Genetics and the Molecular Biology Institute, University of California, Los Angeles, California 90095, USA, [1]Centre de Recherche en Cancerologie de l'Universite Laval, Pavillon H tel-Dieu de Quebec, Quebec, Canada, and [2]Institute of Medical Radiobiology of the University of Zürich and the Paul Scherrer institute, August Forel-Strasse 7, CH-8008 Zürich, Switzerland

*To whom correspondence should be addressed at: Dept. of Microbiology and Molecular Genetics, 1602 Molecular Sciences Bldg., 405 Hilgard Ave., Los Angeles, CA 90095. Tel: 310-825-8460; Fax: 310-206-3088; Email: jhmiller@mbi.ucla.edu

¶Current address: Center for Astrobiology, IGPP, UCLA

‡Current address: Diversa, 4995 Directors Place, San Diego, CA 92121

Running title: Interaction between UDG and PCNA in Crenarchaeon *P. aerophilum*

SUMMARY

Proliferating cell nuclear antigen (PCNA) acts as a sliding clamp on duplex DNA. Its homologs, present in Eukarya and Archaea, are part of protein complexes that are indispensable for DNA replication and DNA repair. In Eukarya, PCNA is known to interact with more than a dozen different proteins, including a human major nuclear uracil-DNA glycosylase (hUNG2) involved in immediate post-replicative repair. In Archaea, only three classes of PCNA binding proteins have been previously reported: replication factor C (the PCNA clamp loader), family B DNA polymerase (Pol B) and flap endonuclease (FEN). In this study we report a direct interaction between a uracil-DNA glycosylase (PaUDGa) and a PCNA homolog (PaPCNA1), both from the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum* (T_{opt} = 100°C). We demonstrate that the PaUDGa-PaPCNA1 complex is thermostable, and two hydrophobic amino acid residues on PaUDGa (Phe^{191} and Leu^{192}) are shown to be crucial for this interaction. It is interesting to note that although PaUDGa has homologs throughout the Archaea and Bacteria, it does not share significant sequence similarity with human major nuclear uracil-DNA glycosylase (hUNG2). Nevertheless, our results raise the possibility that PaUDGa may be a functional analog of hUNG2 for PCNA-dependent post-replicative removal of misincorporated uracil.

INTRODUCTION

Proliferating cell nuclear antigen (PCNA) is essential for life. It is a processivity factor for DNA polymerase, forming a toroidal shaped trimer acting as a sliding clamp on duplex DNA (1-4). Its function requires another protein, the clamp loader replication factor C, to load it onto the circular DNAs (5-9). PCNA is present in eukaryotes and its functional analog, the β subunit of DNA polymerase III holoenzyme, is present in bacteria (10, 11). More than a dozen classes of eukaryotic PCNA binding proteins have been shown to interact with the PCNA sliding clamp, linking PCNA to several important biological processes beyond DNA replication, such as DNA repair and cell cycle regulation (12-18). In many cases PCNA binding partners interact with PCNA through a conserved motif identified as "Qxx(L/M/I)xx(F/Y/H)(F/Y)" that is usually located near either the amino or the carboxyl terminus. One important example of an eukaryotic PCNA binding protein involved in DNA repair is human major nuclear uracil-DNA glycosylase (hUNG2), which removes uracil from misincorporated dUMP residues in an immediate post-replicative process (19, 20). hUNG2 interacts with PCNA through its PCNA binding site, 4-QKTLYSFF-11, which is located near the amino terminus of hUNG2.

Recently PCNA sequence homologs have been identified in Archaea (21, 22). So far, each of the 10 completely sequenced archaeal genomes contains at least one putative PCNA homolog (23). There is a distinction found between the two major subdomains of the Archaea, Crenarchaeota and Euryarchaeota (23). While each euryarchaeal genome tends to have one PCNA homolog, each crenarchaeal genome has two or three putative PCNA homologs (21, 23, 24). Biochemical studies have been conducted with several of the archaeal PCNA homologs, including a PCNA homolog from the euryarchaeote *Pyrococcus furiosus* and two PCNA homologs from the crenarchaeote *Sulfolobus solfataricus* (21, 22). These studies have confirmed that all of them are processivity factors for their corresponding DNA polymerases.

In Archaea, in addition to the PCNA clamp loader (replication factor C), two classes of archaeal proteins

have so far been identified as PCNA binding proteins, based on *in vitro* binding study and crystal structure analysis. They are family B DNA polymerase (Pol B) (22) and flap endonuclease (FEN) (25-29), both proteins known to interact with PCNA in eukaryotes. The proposed putative PCNA binding motifs in these archaeal PCNA binding proteins are quite similar to the conserved PCNA binding motif identified in eukaryotic PCNA binding proteins (22, 27). However, these putative PCNA binding sites have not been verified by mutation analysis.

In this study we report identification of another archaeal PCNA binding protein, *Pyrobaculum aerophilum* uracil-DNA glycosylase 1 (PaUDGa), and the biochemical confirmation of its interaction with PCNA via the PCNA binding motif. *P. aerophilum* is a hyperthermophile with an optimal growth temperature of 100°C, and a member of the crenarchaeal subdomain of Archaea (30). The biochemical characterization of PaUDGa's uracil-DNA glycosylase activity was recently published (31). Analysis of the complete genome sequence of *P. aerophilum* revealed two putative PCNA homologs (24), PaPCNA1 and PaPCNA2, as expected for a crenarchaeote (23). We demonstrate that PaUDGa preferentially binds to PaPCNA1, similar to two other *P. aerophilum* PCNA binding proteins, PaFEN and PaPol B3. PaUDGa's ability to bind to PaPCNA1 resembles the eukaryotic PCNA binding protein hUNG2, which belongs to a distinctly different UDG family due to low amino acid sequence similarity to PaUDGa. Our results raise the possibility that PaUDGa may be a functional analog of hUNG2 for PCNA-dependent post-replicative removal of misincorporated uracil.

EXPERIMENTAL PROCEDURES

Bacterial Expression Plasmids

P. aerophilum genomic DNA was prepared as described previously (32). The coding regions for PaUDGa (PAE0651, accession number AAL62921) and PaFEN (PAE0698, accession number AAL62961) were amplified by polymerase chain reaction (PCR) using *P. aerophilum* genomic DNA as template with their corresponding primer pairs synthesized by Invitrogen (Invitrogen, Carlsbad, CA). The PCR products were cloned into a pCR2.1-TOPO vector using a TOPO TA cloning kit (Invitrogen, Carlsbad, CA). The primer information can be obtained upon request.

The full-length PaUDGa gene was amplified by PCR using pCR2.1TOPOPaUDGa as template, cloned into a pGEX-2TK vector (Amersham Pharmacia Biotech, Piscataway, NJ) at the BamHI site to create a plasmid that expresses a fusion protein of glutathione S-transferase (GST) and PaUDGa. The full-length PaFEN gene was amplified by PCR using pCR2.1TOPOPaFEN as template, cloned into a pGEX-2TK vector (Amersham Pharmacia Biotech) at the EcoRI site to create a plasmid that expresses GST-PaFEN fusion protein. Two PaPol B3 (PAE2109, accession number AAL63952) fragments containing the C-terminal region (C1, amino acid residuals 612-785 and C2, amino acid residuals 726-785) were amplified by PCR using *P. aerophilum* genomic DNA as template, cloned into a pGEX-2TK vector (Amersham Pharmacia Biotech) between BamHI-EcoRI site to create two plasmids that express GST-PaPol B3 (C1, amino acids 612-785) fusion and GST-PaPol B3(C2, amino acids 726-785) fusion, respectively. The full-length PaPCNA1 gene (PAE3038, accession number AAL64629) was amplified by PCR using *P. aerophilum* genomic DNA, cloned into a pQE30 vector (Qiagen, Chatsworth, CA) between BamHI–HindIII site to create a plasmid that expresses the N terminal hexahistidine tagged PaPCNA1. The full-length PaPCNA2 gene (PAE0720, accession number AAL62977) was amplified by PCR using *P. aerophilum* genomic DNA, cloned into a pQE30 vector (Qiagen) between SphI-SalI site to create a plasmid that expresses the N terminal hexahistidine-tagged PaPCNA2. The murine PCNA (33) was subcloned into a pBluescriptII KS vector (Stratagene, La Jolla, CA) between BamHI and EcoRI site. Subsequently the BamHI-HindIII fragment was subcloned into a pQE30 vector (Qiagen) between BamHI and HindIII site to create a plasmid that expresses the N terminal hexahistidine-tagged murine PCNA.

For thermostable binding assays the BamHI-BamHI fragment of PaUDGa was subcloned into pQE30 vector (Qiagen) to create a plasmid (pQE30PaUDGa) that expresses a N-terminal hexahistidine-tagged PaUDGa recombinant protein. The BamHI-HindIII fragment of PaPCNA1 was subcloned into a pQE60 vector (Qiagen) to create a plasmid (pQE60PaPCNA1) that expresses the native form of PaPCNA1 without the histidine tag.

Generation of PaUDGa and PaFEN mutants

The amino acid fragments of PaUDGa 1-182, 131-196, and 172-196 were amplified by PCR using pCR2.1TOPOPaUDGa as template. The products were subcloned into a pGEX-2TK vector (Amersham Pharmacia Biotech) at BamHI site to create GST fusion protein expression plasmids.

The PaUDGa mutant F183A/F184A was generated by standard site-directed mutagenesis procedure (34) using pCR2.1TOPOPaUDGa as template. The obtained PCR products were cloned into a pGEX-2TK vector (Amersham Pharmacia Biotech) at BamHI site to create an expression plasmid for GST–PaUDGa (F183A/F184A) and the mutations were verified by DNA sequencing using a SequiTherm EXCEL II DNA sequencing kit (Epicentre, Madison, WI).

The PaUDGa mutant F191A/L192A was generated by site-directed mutagenesis. The PCR products were first cloned into a pCR2.1-TOPO vector using a TOPO TA cloning kit (Invitrogen). The BamHI-BamHI fragment containing the full-length PaUDGa with two amino acid substitutions was subcloned into a pGEX-2TK vector (Amersham Pharmacia Biotech) at BamHI site to create an expression plasmid for GST-PaUDGa (F191A/L192A) and the mutations were verified by DNA sequencing using a SequiTherm EXCEL II DNA sequencing kit (Epicentre).

The PaFEN mutant F345A/F346A was generated by site-directed mutagenesis. The PCR products were first cloned into a pCR2.1-TOPO vector using a TOPO TA cloning kit (Invitrogen). The EcoRI-EcoRI fragment containing the full-length PaFEN with two amino acid substitutions was subcloned into a pGEX-2TK vector (Amersham Pharmacia Biotech) at EcoRI site to create an expression plasmid for GST-PaFEN (F345A/F346A) and the mutations were verified by DNA sequencing using a SequiTherm EXCEL II DNA sequencing kit (Epicentre).

Expression and partial purification of recombinant PCNA homologs.

Overnight cultures of *E. coli* BL21/pREP4 harboring plasmid pQE30-PaPCNA1,

pQE30-PaPCNA2, or pQE30-murinePCNA were used to inoculate 100 ml of LB medium supplemented with 200 μg/ml Ampicillin and 25 μg/ml kanamycin. The expression of PaPCNA1, PaPCNA2, or murine PCNA was induced with 0.3 mM isopropyl-β-D-thiogalactopyranoside (IPTG) during mid-log growth phase for 3 hr at 37°C. The plasmid pREP4 constitutively expresses the Lac repressor protein encoded by the *lac*I gene in order to reduce the basal level of expression (Qiagen). Cells were lysed in buffer A (20 mM Tris-HCl, pH 7.4, 60 mM NaCl, and 2 mM EDTA, supplemented with 0.2 mg/ml lysozyme and 1 mM PMSF) by sonication. The cell lysates were clarified by centrifugation at 15,000 x g for 30 min. The protein concentration was determined by Bradford Protein Assay (BioRad). The cell lysates were stored as small aliquots at –80°C. Partial purification of PaPCNA1 and PaPCNA2 was obtained by heat treatment of 200 μl cell lysates at 70°C for 10 min followed by centrifugation at 15,000 x g for 30 min. The supernatant was stored at 80°C.

A native PaPCNA1 recombinant protein without the histidine tag was obtained for the thermostable binding assay by a similar protocol as described above. An *E. coli* strain BL21/pREP4 harboring plasmid pQE60PaPCNA1 was used for the preparation of cell lysates. Subsequently, a heat treatment procedure was employed to the cell lysates to obtain a partial purified recombinant PaPCNA1 in its native form without the histidine tag.

Expression, immobilization of GST fusion proteins, and "pull-down" affinity bead interaction assay

E. coli BL21 strain was used as the host to express all GST fusion proteins and the GST protein. The cell lysates were prepared in buffer A as described in the above section. Protein concentration of the cell lysate was determined by Bradford protein assay (BioRad). Cell lysates were incubated with the affinity matrices (GST sepharose beads, Amersham Pharmacia Biotech) for 1 h at 4°C. After extensive washing with buffer A, GST

sepharose beads with immobilized GST fusion proteins were stored in buffer A at 4°C.

To determine the amount of each GST fusion protein bound for the "pull-down" assay, a fraction of GST sepharose bead-immobilized proteins were released by boiling in SDS sample buffer, analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), and visualized by Coomassie Brilliant blue staining.

The amount of each GST fusion protein bound used for the pull-down assay was adjusted for each sample based on the amount of bound proteins analyzed by SDS-PAGE. Extra GST sepharose beads were added to samples to bring the final amount of 60 µl. GST sepharose beads were mixed with PCNA samples at 4°C for 1h. After six-time washing with 0.8 ml buffer A, the bound proteins were released by boiling in 60 µl SDS sample buffer, analyzed by SDS-PAGE and transferred to nitrocellulose membrane. Western blot analysis for PaPCNA homologs were performed using Manufacturer's protocols (Qiagen) with mouse anti-RGS$(H)_4$ primary antibody (1:500) and HRP-conjugated secondary antibody (1:5000) followed by ECL detection (Amersham Pharmacia Biotech).

Thermostable binding assay

A N-terminal hexahistidine-tagged PaUDGa protein was expressed in an *E. coli* BL21/pREP4/pQE30PaUDGa strain and the cell lysate was prepared by sonication in buffer B (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, supplemented with 1 mM PMSF). To immobilize PaUDGa, Ni^{2+}-NTA beads were added to the cell lysate and incubated at 4°C for 1 h. After extensive washing with the buffer B, the Ni^{2+}-NTA beads were transferred to buffer A. The Ni^{2+}-NTA beads (60 µl) containing immobilized PaUDGa were mixed with the partial purified recombinant PaPCNA1 in its native form or with *E. coli* cell lysate containing pQE30 vector alone after 70°C heat-treatment. As a

control for the experiment, a sample containing the partial purified PaPCNA1 and 60 µl Ni^{2+}-NTA beads without immobilizes PaUDGa was also prepared. After five-time washing with 0.8 ml buffer A at 4°C, the Ni^{2+}-NTA beads were evenly split into two tubes. While one tube was incubated at 70°C waterbath for 5 min with 0.8 ml buffer A pre-equilibrated at 70°C, the other tube was kept on ice with 0.8 ml chilled buffer A. Ni^{2+}-NTA beads were pelleted at 1000 x g for 2 min. Bound proteins were released by boiling in SDS sample buffer, analyzed by SDS-PAGE and visualized by Coomassie Brilliant blue staining.

RESULTS

Expression and partial purification of two putative PCNA homologs from *P. aerophilum*

As a member of the archaeal subdomain Crenarchaeota, *P. aerophilum* contains two putative PCNA homologs, PaPCNA1 and PaPCNA2, identified by amino acid sequence homology (Fig. 1A, ref. 24). PaPCNA1 shares 24% amino acid sequence identity with PaPCNA2 (Fig 1B). The amino acid sequence identities between the two putative PaPCNA homologs and the other characterized eukaryotic or archaeal PCNAs range from 17% to 26% (Fig. 1B). Both putative PaPCNA homologs contain the highly conserved (L/I)-A-P-(K/R) motif located near the carboxyl terminus, which may interact with the clamp loader, replication factor C (23). Phylogenetic analysis of PCNA homologs in crenarchaea reveals that the multiple homologs in Crenarchaea fall into two classes, consistent with a duplication event early after the divergence of the crenarchaeal clade (23).

PaPCNA1 and PaPCNA2 were cloned on the bacterial expression vector pQE30 and expressed as N-terminal hexahistidine-tagged recombinant proteins in *E. coli* (Fig. 2A and 2B,

lanes 3 and 5). To obtain partially purified recombinant PaPCNA homologs for *in vitro* binding assays, we took advantage of the heat stability characteristic of proteins from thermophiles and heated the *E. coli* crude cell lysates expressing each homolog to 70°C. The recombinant PaPCNA1 and PaPCNA2 were largely heat-stable after 10-min heat treatment (Fig. 2A and 2B, lanes 4 and 6). A hexahistidine-tagged eukaryotic murine PCNA recombinant protein was also included in the experiment and was found to be heat labile under the conditions tested (Fig. 2A and 2B, lanes 7 and 8). Although the recombinant PaPCNA1 and PaPCNA2 have predicted molecular masses of 28 and 29 kDa, respectively, the apparent molecular mass on the SDS-polyacrylamide gel was approximately 37 kDa for PaPCNA1 and 32 kDa for PaPCNA2. The aberrant migration of PaPCNAs on the SDS-polyacrlamide gel was also seen in two *S. solfataricus* PCNAs (21) and eukaryotic PCNA homologs (Fig. 2A, lane 7, ref. 35) for unknown reasons.

***In vitro* direct binding of PaPol B3 to PaPCNA1**

P. aerophilum DNA polymerase B3 (PaPol B3) was tested for binding to recombinant PaPCNA1 and PaPCNA2. Analysis of the PaPol B3 protein sequence revealed that its C terminal region contains a putative PCNA binding motif "778-ERTLLDFF-786" (Fig. 3A), which is similar to the putative PCNA binding motifs of several archaeal Pol B homologs predicted by Ishino and his colleagues (22). Glutathione S-transferase (GST) fusion proteins were constructed for two overlapping fragments containing the carboxyl terminal region of PaPol B3 (C1↑amino acids 612-785, C2↑amino acids 726-785, Figs 3A & 3B). In the "pull-down" affinity bead interaction assays, GST alone had no detectable binding to either of the PaPCNAs (Fig. 3C, lanes 2 and 6). However, binding to PaPCNA1 was detected with both of the GST-

PaPol B3 fusions (Fig. 3C, lanes 3 and 4). In addition, a weak binding to PaPCNA2 was also detected (Fig. 3C, lanes 7 and 8). The observed preferred binding to PaPCNA1 by the C terminal region of PaPol B3 provides evidence for an *in vivo* direct interaction between PaPol B3 and PaPCNA1.

***In vitro* direct binding of PaUDGa and PaFEN to PaPCNA1**

A GST fusion protein with PaUDGa was constructed for the study of *in vitro* direct binding of PaUDGa to two recombinant PaPCNA homologs using the pull-down affinity bead interaction assay (Fig. 4A, lane 3). PaFEN, which was predicted to be a PCNA binding protein based on previous studies (27), was also included for the binding experiment (Fig. 4A, lane 4). While GST alone bound to neither PaPCNA (Fig. 4B, lanes 3 and 4), both GST-PaUDGa and GST-PaFEN bound to PaPCNA1 (Fig. 4B, lanes 5 and 7). Binding to PaPCNA2 was not detected with either GST fusion (Fig. 4B, lanes 6 and 8). Therefore, the observed binding of PaUDGa and PaFEN to PaPCNA1 provides evidence for a direct *in vivo* interaction between PaUDGa-PaPCNA1 and PaFEN-PaPCNA1. The effect of NaCl concentration on the formation of a complex between PaUDGa and PaPCNA1 was studied. Binding was detected in the presence of 0.05 – 0.4 M NaCl and was disrupted at NaCl concentrations higher than 0.4 M (Fig. 4C and 4D). The effect of temperature on the formation of the complex was also studied. When using a recombinant hexahistidine-tagged PaUDGa and a recombinant PaPCNA1 in its native form without any tag, binding was observed at 4°C and was largely retained after 5 min-treatment at 70°C (Fig. 4E, lanes 3 and 4). Under the same experimental conditions, the recombinant PaPCNA1 protein alone had no detectable binding to the Ni^{2+}-NTA resin (data not shown). Binding of PaUDGa and PaFEN to the eukaryotic murine PCNA was also tested, but

was undetectable under our experimental conditions (data not shown).

The PCNA interaction motif is located near the C-terminus of PaUDGa

The interaction between PaUDGa and PaPCNA1 was further verified by identifying the specific regions on PaUDGa required for PCNA binding. First, three GST fusion proteins that contain various regions of PaUDGa were constructed and tested for PaPCNA1 binding activity using the pull-down affinity bead interaction assay (Fig. 5A). GST-PaUDGa (amino acids 1-182) did not bind to PaPCNA1 and this fusion protein had an excessive proteolytic degradation (data not shown). However, both GST fusion proteins containing the carboxyl terminal region of PaUDGa displayed binding activity (data not shown for GST-PaUDGa (amino acids 131-196); see Fig. 5C and 5D, lane 8 for GST-PaUDGa (amino acids 172-196)). These results demonstrate that the 25-amino acid region near the PaPCNA1 carboxyl terminus is required for PCNA binding.

Next, PaUDGa mutants with specific substitutions within the above identified 25 amino acid region were generated in order to identify amino acid residues critical for PCNA binding activity. Previous studies have shown that many PCNA binding proteins contain the consensus PCNA binding motif "Qxx(L/M/I)xx(F/Y/H)(F/Y)" and mutational analysis indicates that the two consecutive hydrophobic amino acid residuals within this motif are involved in the interaction with PCNA (36). Analysis of the PaUDGa amino acid sequence revealed two putative PCNA binding motifs near the carboxyl terminus of PaUDGa (Fig. 5B). The sequence for the first region is 177-QKDLAMFF-184 (motif 1), which contains the eukaryotic PCNA binding consensus sequence "QxxLxxFF". The sequence for the second region is 185-GGGLDRFL-192 (motif 2), which contains two consecutive hydrophobic amino acid residues (Phe^{191} and Leu^{192}) closer to the carboxyl terminus (Fig. 5B). Two PaUDGa mutants were generated for the pull-down binding assay, each with two amino acid changes in the putative PCNA binding motif 1 (F183A/F184A) or motif 2 (F191A/L192A). A PaFEN mutant carrying F345A/F346A modification in its putative PCNA binding motif 339-TSSLDSFF-346 near its carboxyl terminus was also included in the experiment (Fig.

5B). As expected, the PaFEN mutant carrying F345A/F346A failed to bind to PaPCNA1 (Fig. 5C and 5D, lane 4). The PaUDGa mutant carrying F183A/F184A in the putative binding motif 1 was still capable of binding to PaPCNA1 (Fig. 5C and 5D, lane 6). However, the binding was largely abolished in the PaUDGa mutant carrying F191A/L192A in the putative motif 2 (Fig. 5C and 5D, lane 7). These results show that Phe345 and Phe346 of PaFEN, and Phe191 and Leu192 of PaUDGa are necessary for the binding of PaFEN and PaUDGa to PaPCNA1.

DISCUSSION

With the goal of finding new archaeal PCNA binding proteins, we searched the predicted protein sequences of the *P. aerophilum* genome with identified putative PCNA binding motifs using MACPATTERN 3.6 (37). In this way PaUDGa was identified as a putative PCNA binding protein. We carried out *in vitro* biochemical analysis of PaUDGa binding to two putative *P. aerophilum* PCNA homologs, PaPCNA1 and PaPCNA2, using the pull-down affinity bead interaction assay. Our results show that PaUDGa preferentially binds to PaPCNA1 to form a thermostable protein complex. Apparently, the binding between PaUDGa and PaPCNA1 is specific, as PaUDGa has no detectable binding to the murine PCNA. Two consecutive hydrophobic amino acid residues Phe^{191} and Leu^{192}, located near the carboxyl terminus of PaUDGa, are crucial for PCNA binding activity. Maintaining these same experimental conditions, the interactions between PaPCNA1 and two other *P. aerophilum* proteins–PaFEN (wild type and mutant F345A/F346A) and the C terminal region of PaPol B3–provides evidence for PaPCNA1 as a functional PCNA homolog in *P. aerophilum*. At this point we were unable to perform *in vivo* experiments to verify the physiological significance of the PCNA binding property of PaUDGa due to the current lack of a genetic system in *P. aerophilum*. However, identification of PaUDGa as an archaeal PCNA binding protein and its critical amino acid residues for PCNA binding is the first step in revealing its biological significance in future

genetic analyses.

The PCNA binding activity observed in PaUDGa from this study and in hUNG2 from previous studies (19) raises the possibility that PaUDGa may be a functional analog of hUNG2. PaUDGa and hUNG2 belong to two distinct UDG families due to the low sequence similarity between them (31, 38). In addition, PaUDGa is an olive green colored protein (31) due to the presence of an iron-sulfur cluster (L. Pearl, personal communication), while hUNG2 is largely colorless. Despite these differences, both PaUDGa and hUNG2 have similar uracil glycosylase activities and both interact with PCNA through specific binding domains at either the carboxyl terminal region (for PaUDGa) or the amino terminal region (for hUNG2, 19). Thus, it is possible that PaUDGa may be a functional analog of hUNG2 for PCNA-dependent post-replicative removal of misincorporated uracil (19, 20).

At this point, two or three putative PCNA homologs have been identified in the completely sequenced Crenarchaeota genomes, including *P. aerophilum*, S. *solfataricus*, and *A. pernix* (21, 23). The presence of more than one PCNA homolog may reflect either functional redundancy or functional differentiation. Previous experiments with *S. solfataricus* SsoPCNA A (039p) and SsoPCNA B (048p) demonstrate that both SsoPCNA homologs are processivity factors for the single-subunit family B DNA polymerase (Pol B1) despite their slightly different efficiencies (21). Our study with PaPCNA1 and PaPCNA2 suggests a possible functional differentiation between two PaPCNA homologs. Binding to PaUDGa and PaFEN is only detected with PaPCNA1. Similarly, the carboxyl terminal region of PaPol B3 also preferentially binds to PaPCNA1. Although we cannot eliminate the possibility that the recombinant PaPCNA2 is somehow inactivated under our experimental conditions–regardless of its heat

stability and weak binding to the C terminal region of PaPol B3–our results suggest that PaPCNA1 may be the major protein with functions similar to PCNA in eukaryotes. Whether PaPCNA2 is functional and if so, which proteins it might interact with remains to be answered. Our experiments do not rule out the possibility that PaPCNA1 and PaPCNA2 monomers may combine to make functional heteromeric trimers.

While FEN homologs are present in Archaea and Eukarya (27), PaUDGa family members are found in Archaea and Eubacteria (38, 39). The amino acid sequence alignments of the carboxyl terminal regions of FEN and UDG homologs are shown in Figure 6. While the conserved PCNA binding motif "QxxLxx(F/W)F" was observed near the carboxyl terminus for almost all archaeal FEN homologs (Fig. 6A, ref. 27), divergent carboxyl terminal sequences were observed for archaeal UDG homologs (Fig. 6B). Only four homologs contained recognizable PCNA binding motifs: ApeUDG from *Aeropyrum pernix*, SsoUDG1 from *S. solfataricus*, AfUDG from *Archaeoglobus fulgidus*, and HaloUDG from *Halobacterium* sp. NRC-1. Biochemical experiments will be required to test these binding motifs and to determine whether the remaining archaeal homologs also bind PCNA.

ACKNOWLEDGEMENT

This work was supported by USHHS Institutional National Research Service Award

#T32 CA09056 (to H.Y.) and by National Institute of Health Grant GM 57917 (to J.H.M.). The generous financial support of the UBS Stiftung (to A.A.S) is gratefully acknowledged.

REFERENCES

1. Kong, X. P., Onrust, R., O'Donnell, M., and Kuriyan, J. (1992) *Cell* **69**, 425-437
2. Krishna, T. S., Kong, X. P., Gary, S., Burgers, P. M., and Kuriyan, J. (1994) *Cell* **79**, 1233-1243
3. Moarefi, I., Jeruzalmi, D., Turner, J., O'Donnell, M., and Kuriyan, J. (2000) *J. Mol. Biol.* **296**, 1215-1223
4. Matsumiya, S., Ishino, Y., and Morikawa, K. (2001) *Protein Sci.* **10**, 17-23
5. Pisani, F.M., De Felice, M., Carpentieri, F., and Rossi, M. (2000) *J. Mol. Biol.* **301**, 61-73
6. Oyama, T., Ishino, Y., Cann, I.K., Ishino, S., and Morikawa, K. (2001) *Mol. Cell* **8**, 455-463
7. Jeruzalmi, D., Yurieva, O., Zhao, Y., Young, M., Stewart, J., Hingorani, M., O'Donnell, M., and Kuriyan, J. (2001a) *Cell* **106**, 417-428
8. Jeruzalmi, D., O'Donnell, M., and Kuriyan, J. (2001b) *Cell* **106**, 429-441
9. Cann, I.K., Ishino, S., Yuasa, M., Daiyasu, H., Toh, H., and Ishino, Y. (2001) *J. Bacteriol.* **183**, 2614-2623
10. Hingorani, M. M., and O'Donnell, M. (2000) *Curr Biol.* **10**, R25-29
11. Bruck, I., and O'Donnell, M. (2001) *Genome Biol.* **2**, review3001.1-3001.3
12. Kelman, Z., and Hurwitz, J. (1998) *Trends Biochem Sci.* **23**, 236-238
13. Tsurimoto, T. (1999) *Front Biosci.* **4**, D849-858
14. Krude, T. (1999) *Curr Biol*. **9**, R394-396
15. Kolodner, R.D., and Marsischky, G.T. (1999) *Curr Opin Genet Dev*. **9**, 89-96
16. Warbrick, E. (2000) The puzzle of PCNA's many partners. *Bioessays* **22**, 997-1006
17. Gulbis, J.M., Kelman, Z., Hurwitz, J., O'Donnell, M., and Kuriyan, J. (1996) *Cell* **87**, 297-

306

18. Shamoo, Y., and Steitz, T.A. (1999) *Cell* **99**, 155-166

19. Otterlei, M., Warbrick, E., Nagelhus, T.A., Haug, T., Slupphaug, G., Akbari, M., Aas, P.A., Steinsbekk, K., Bakke, O., and Krokan, H.E. (1999) *EMBO J.* **18**, 3834-3844

20. Krokan, H.E., Otterlei, M., Nilsen, H., Kavli, B., Skorpen, F., Andersen, S., Skjelbred, C., Akbari, M., Aas, P.A., and Slupphaug, G. (2001) *Prog Nucleic Acid Res Mol Biol.* **68**, 365-386

21. De Felice, M., Sensen, C. W., Charlebois, R. L., Rossi, M., and Pisani, F. M. (1999) *J. Mol. Biol.* **291**, 47-57

22. Cann, I.K., Ishino, S., Hayashi, I., Komori, K., Toh, H., Morikawa, K. and Ishino, Y. (1999) *J. Bacteriol.* **181**, 6591-6599

23. Iwai, T., Kurosawa, N., Itoh, Y. H., and Horiuchi, T. (2000) *Extremophiles* **4**, 357-364

24. Fitz-Gibbon, S.T., Ladner, H., Kim, U.J., Stetter, K.O., Simon, M.I., and Miller, J.H. (2002) *Proc. Natl. Acad. Sci.* U S A. **99**, 984-989

25. Hwang, K.Y., Baek, K., Kim, H.Y., and Cho, Y. (1998) *Nat Struct Biol.* **5**, 668-670

26. Hosfield, D.J., Mol, C.D., Shen, B., and Tainer, J.A. (1998a) *Cell* **95**, 135-146

27. Hosfield, D.J., Frank, G., Weng, Y., Tainer, J.A., and Shen, B. (1998b) *J. Biol. Chem*. **273**, 27154-27161

28. Rao, H.G., Rosenfeld, A., and Wetmur, J.G. (1998) *J. Bacteriol*. **180**, 5406-5412

29. Matsui, E., Kawasaki, S., Ishida, H., Ishikawa, K., Kosugi, Y., Kikuchi, H., Kawarabayashi, Y., and Matsui, I. (1999) *J. Biol. Chem*. **274**, 18297-18309

30. Völkl, P., Huber, R., Drobner, E., Rachel, R., Burggraf, S., Trincone, A., and Stetter, K.O.

(1993) *Applied and Environmental Microbiol.* **59**, 2918-2926

31. Sartori, A.A., Schar, P., Fitz-Gibbon, S., Miller, J.H., and Jiricny, J. (2001) *J. Biol. Chem.* **276**, 29979-29986

32. Fitz-Gibbon, S., Choi, A.J., Miller, J.H., Stetter, K.O., Simon, M.I., Swanson, R., and Kim, U.J. (1997) *Extremophiles* **1**, 36-51

33. Lebel, M., Spillare, E.A., Harris, C.C., and Leder, P. (1999) *J. Biol. Chem*. **274**, 37795-37799

34. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor, NY

35. Zhang, P., Zhang, S.-J., Zhang, Z., Woessner, Jr. J.F., and Lee, M.Y.W.T. (1995) *Biochemistry* **34**, 10703-10712

36. Haracska, L., Johnson, R.E., Unk, I., Phillips, B., Hurwitz, J., Prakash, L., and Prakash, S. (2001) *Mol. Cell Biol*. **21**, 7199-7206

37. Fuchs, R. (1994) *Computer Applications in the Biosciences* **10**, 171-178

38. Sandigursky, M., and Franklin, W.A. (1999) *Curr. Biol*. **9**, 531-534

39. Sandigursky, M., and Franklin, W.A. (2000) *J. Biol. Chem*. **275**, 19146-19149

40. Thompson, J.D., Higgins, D.G. and Gibson,T.J. (1994) *Nucleic Acids Res.,* **22**, 4673–4680

FIGURE LEGENDS

Figure 1. Sequence alignment of PaPCNA1, PaPCNA2, and characterized eukaryotic and archaeal PCNA homologs using the program ClustalW (40) (A) and percentage of sequence identity among PCNA homologs shown in A (B). The homologs are the following (organisms and the protein accession numbers in parentheses): HsaPCNA (*Homo sapiens*, P12004); SpoPCNA (*Schizosaccharomyces pombe*, CAB38513); PfuPCNA (*P. furiosus*, BAA33020); PaPCNA1 (*P. aerophilum* PCNA1, AAL64629); SsoPCNA A (039p) (*S. solfataricus*, P57765); PaPCNA2 (*P. aerophilum*, AAL62977); SsoPCNA B (048p) (*S. solfataricus*, P57766). Black boxes indicate 70-100% amino acid residue identity, gray ones indicate 70-100% similarity. The asterisks mark amino acid residues corresponding to those forming the hydrophobic pocket in human PCNA (17, 23).

Figure 2. Partial purification of recombinant PaPCNA1 and PaPCNA2. The arrows mark full-length recombinant proteins. (A) SDS-PAGE analysis of crude *E. coli* lysates containing hexahistidine-tagged PaPCNA1 and hexahistidine-tagged PaPCNA2 before the heat treatment (4°C) and after the heat treatment (70°C). Vector - pQE30 without insert (lanes 1 and 2); PaPCNA1 - pQE30PaPCNA1 (lanes 3 and 4); PaPCNA2 - pQE30PaPCNA2 (lanes 5 and 6); and mPCNA - pQE30murinePCNA (expresses hexahistidine-tagged murine PCNA, lanes 7 and 8). The crude *E. coli* lysates were prepared as described in Materials and Methods, subjected to SDS-PAGE and visualized by Coomassie Blue staining. (B) Western blot analysis of the SDS gel shown in A. The Western blotting procedure was performed as described in Material and Methods.

Figure 3. *In vitro* direct binding between PaPCNA1 and the C terminal region of PaPol B3. (A) Schematic diagram of PaPol B3 and the two fragments (C1 and C2). (B) SDS-PAGE analysis of the following samples: GST (lane 2), GST-Pol B3 (C1, amino acids 612-785) (lane 3) and GST-Pol B3 (C2, amino acids 726-785) (lane 4). Lane M contains molecular mass standards (Bio-Rad) as indicated on the left. The protein bands were visualized by Coomassie Blue staining. (C) Western blot analysis of the samples from the pull-down experiment: inputs of PaPCNA1 and PaPCNA2 (lanes 1 and 5); GST alone with PaPCNA1 and PaPCNA2 (lanes 2 and 6); GST-Pol B3(C1) with PaPCNA1 and PaPCNA2 (lanes 3 and 7); GST-Pol B3(C2) with PaPCNA1 and PaPCNA2 (lanes 4 and 8). Sample preparation and western blot analysis are described in Material and Methods.

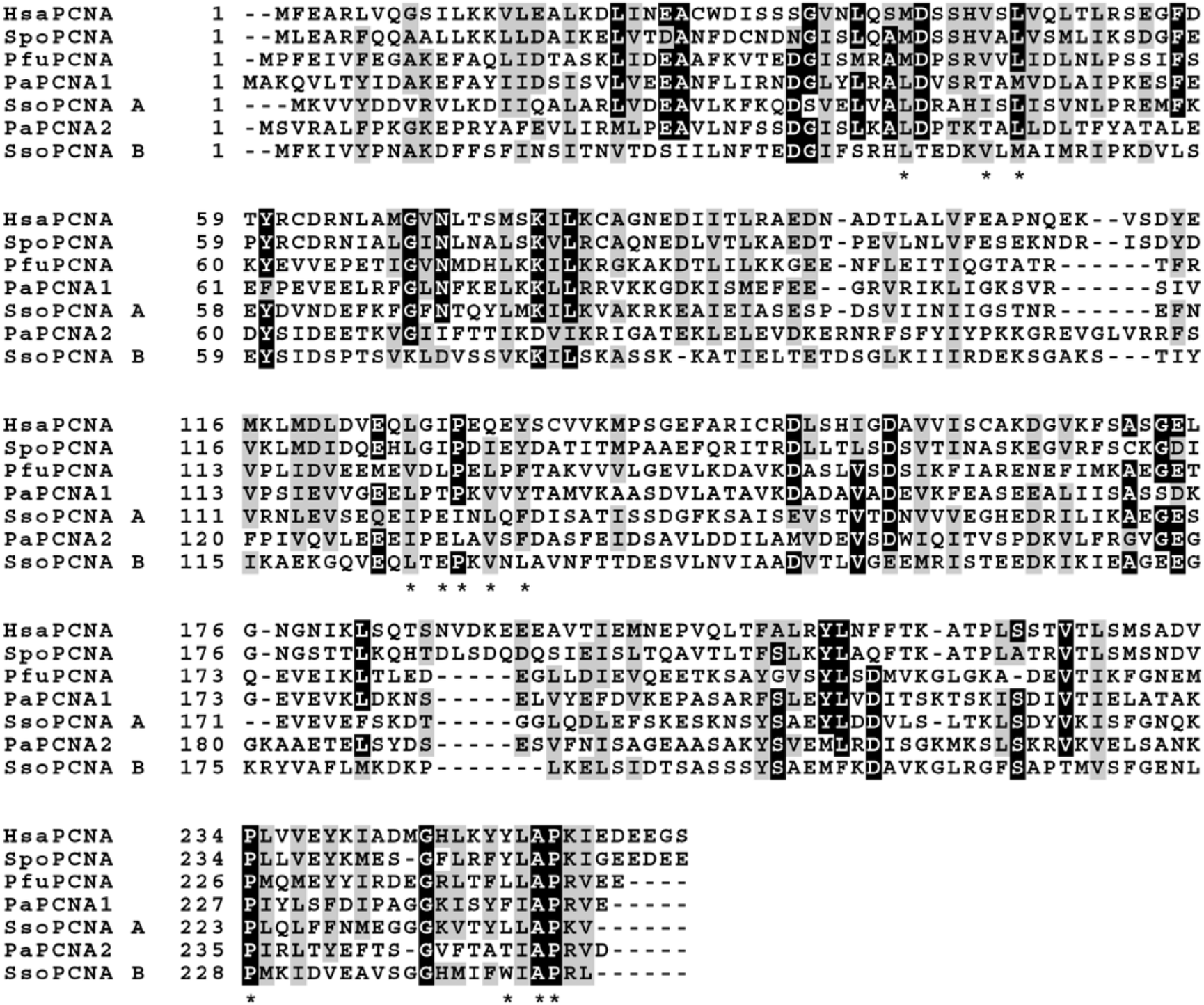
Figure 4. *In vitro* direct binding of PaUDGa and PaFEN to PaPCNA1, the effect of NaCl concentration and temperature. (A) SDS-PAGE analysis of the following samples: GST (lane 2), GST-PaUDGa (lane 3) and GST-PaFEN (lane 4). Lane M contains molecular mass standards (Bio-Rad) as indicated on the left. The protein bands were visualized by Coomassie Blue staining. (B) Western blot analysis of the samples from the pull-down experiment: inputs of PaPCNA1 and PaPCNA2 (lanes 1 and 2); GST alone with PaPCNA1 and PaPCNA2 (lanes 3 and 4); GST-PaUDGa with PaPCNA1 and PaPCNA2 (lanes 5 and 6); GST-PaFEN with PaPCNA1 and PaPCNA2 (lanes 7 and 8). (C) Western blot analysis of the samples from the pull-down experiment of GST-PaUDGa with PaPCNA1 in the presence of varying amounts of NaCl as indicated. The position of PaPCNA1 was marked on the right. (D) SDS-PAGE analysis of the amount of full-length GST-PaUDG fusion protein used in the pull-down

experiment shown in C. The protein bands were visualized by Coomassie blue staining. The position of the full-length GST-PaUDGa was marked on the right. (E) Thermostable binding between PaUDGa and PaPCNA1. SDS-PAGE of the following samples from the pull-down experiment: input of PaPCNA1 without any tag (lane 2); 6His-tagged PaUDGa with PaPCNA1 without 70°C treatment (lane 3) and with 70°C treatment (lane 4). The protein bands were visualized by Coomassie blue staining. Lane M contains molecular mass standards (Bio-Rad) as indicated on the left. Sample preparation and western blot analysis are described in Material and Methods.

Figure 5. PCNA binding activity of the C terminal region of PaUDGa. (A) Schematic diagram summarizing the results of the PCNA binding assays with GST fusions containing different PaUDGa amino acid segments as indicated. (B) Schematic diagrams of PaUDGa and PaFEN proteins. The putative PCNA binding motifs located at the C terminal region of PaUDGa and PaFEN are indicated. (C) SDS-PAGE analysis and (D) western blot analysis of the following samples from the pull-down experiment: input of PaPCNA1 (lane 1); GST alone with PaPCNA1 (lane 2); GST-wild type PaFEN (WT, lane 3) and GST-mutant PaFEN F345A/F346A (FF→AA, lane 4) with PaPCNA1; GST-wild-type PaUDGa (WT, lane 5) and three GST-PaUDGa mutants with PaPCNA1: F183A/F184A (FF→AA, lane 6); F191A/L192A (FL→AA, lane 7), and C-terminal (amino acids 172-196, lane 8). The position of PaPCNA1 was marked on the right. The molecular mass standards (Bio-Rad) are indicated on the left. Sample preparation and western blot analysis are described in Material and Methods.

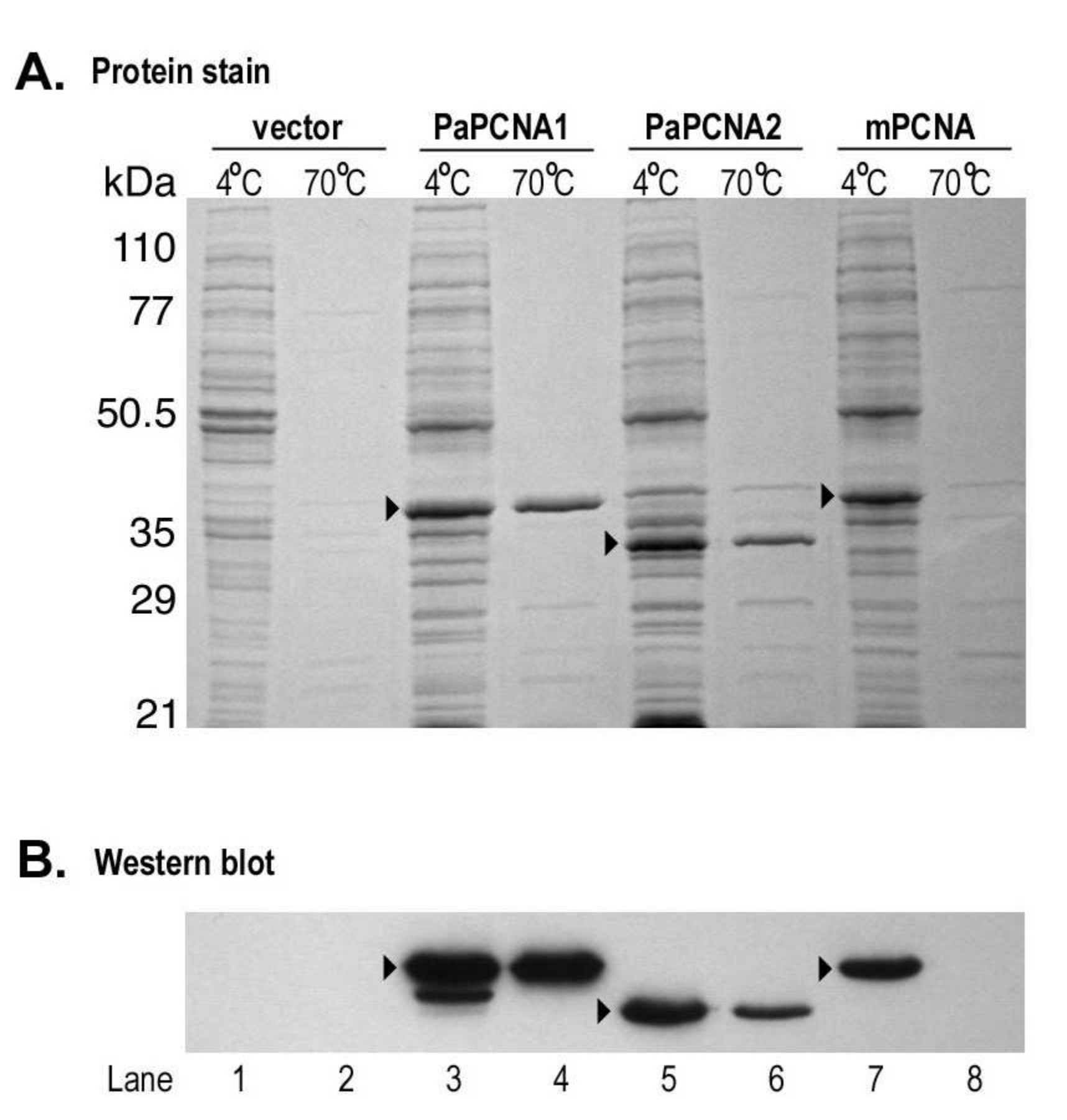
Figure 6. Sequence alignment of the C terminal regions of archaeal FEN homologs and archaeal UDG homologs using the ClustalW program (40). Black boxes indicates 70-100% amino acid identity, gray ones 70-100% amino acid similarity. (A) C terminal region of FEN. The asterisks mark the proposed PCNA binding motif (27). The symbol (u) marks an eukaryotic FEN homolog. The homologs are the following (organisms and the protein accession numbers in parentheses): TvFEN (*Thermoplasma volcanium*, BAB59701; TaFEN (*Thermoplasma acidophilum*, CAC12164); ApeFEN (*Aeropyrum pernix*, BAA79026); SsoFEN (*S. solfataricus*, AAK40525); PaFEN (*P. aerophilum*, AAL62961); AfFEN (*Archaeoglobus fulgidus*, AAB90967); PabFEN (*Pyrococcus abyssi*, CAB49654); PhFEN (*Pyrococcus horikoshii*, BAA30521); MjFEN (*Methanococcus jannaschii*, AAB99454); MthFEN (*Methanobacterium thermoautotrophicum*, AAB86106); HaloFEN (*Halobacterium* sp. NRC-1, AAG19690); and HsaFEN (*Homo sapians*, AAH00323). (B) C-terminal region of UDG. The boxed amino acid residues are proposed to be involved in PCNA interaction. The symbol (s) indicates a eubacterial UDG homolog. The homologs are the following (organisms and the protein accession numbers in parentheses): PaUDGa (*P. aerophilum*, AAL62921); ApeUDG (*A. pernix*, BAA79385); SsoUDG1 (*S. solfataricus*, AAK42437); AfUDG (*A. fulgidus*, AAB88977); PabUDG (*P. abyssi*, CAB49606); PhUDG (*P. horikoshii*, BAA30579); AaUDG (*Aquifex aeolicus*, AAC07559); TmUDG (*Thermotoga maritima*, AAD35596); TpUDG (*Treponema pallidum*, AAC65215); HaloUDG (*Halobacterium* sp. NRC-1, AAG20230); TaUDG (*T. acidophilum*, CAC11619); and TvUDG (*T. volcanium*, BAB59983).

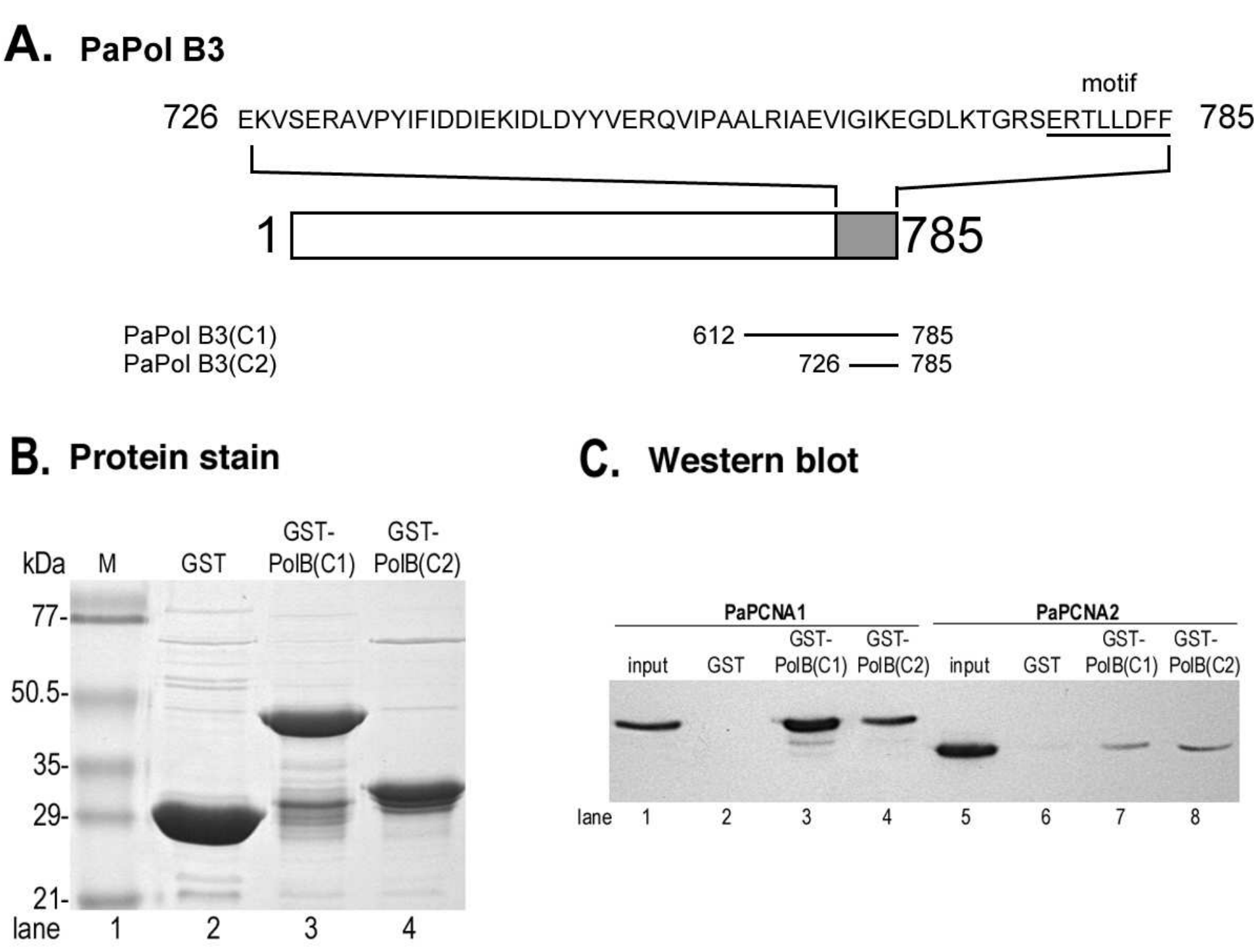
A.



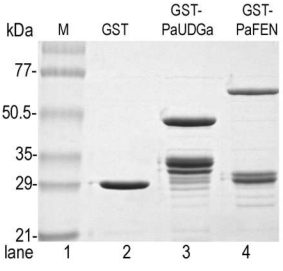
B.

PaPCNA1	24	23	17	25	18	18
PaPCNA2	-	26	24	25	21	19
	PaPCNA2	SsoPCNA A	SsoPCNA B	PfuPCNA	SpoPCNA	HsaPCNA

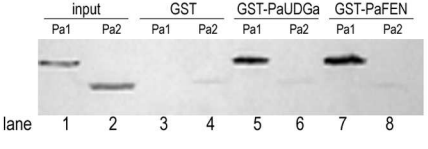

A.
Protein stain
vector
PaPCNA1
PaPCNA2
mPCNA
kDa
4°C
70°C
4°C
70°C
4°C
70°C
4°C
70°C
110
77
50.5
35
29
21
B.
Western blot
Lane
1
2
3
4
5
6
7
8


A. PaPol B3
motif
726 EKVSERAVPYIFIDDIEKIDLDYYVERQVIPAALRIAEVIGIKEGDLKTGRSERTLLDFF 785
1
785
PaPol B3(C1)
PaPol B3(C2)
612 785
726 785
B. Protein stain
kDa
M
GST
GST-PolB(C1)
GST-PolB(C2)
77-
50.5-
35-
29-
21-
lane 1 2 3 4
C. Western blot
PaPCNA1
PaPCNA2
input
GST
GST-PolB(C1)
GST-PolB(C2)
input
GST
GST-PolB(C1)
GST-PolB(C2)
lane 1 2 3 4 5 6 7 8

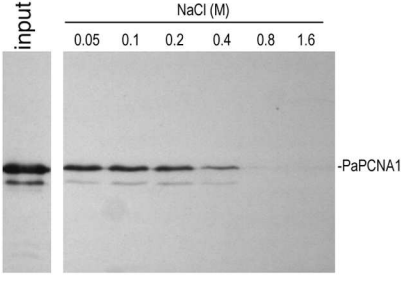
A. Protein stain



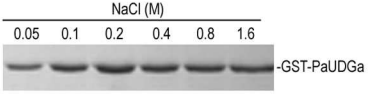
B. Western blot



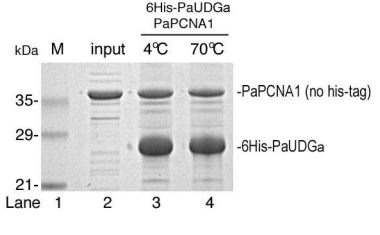
C. Western blot

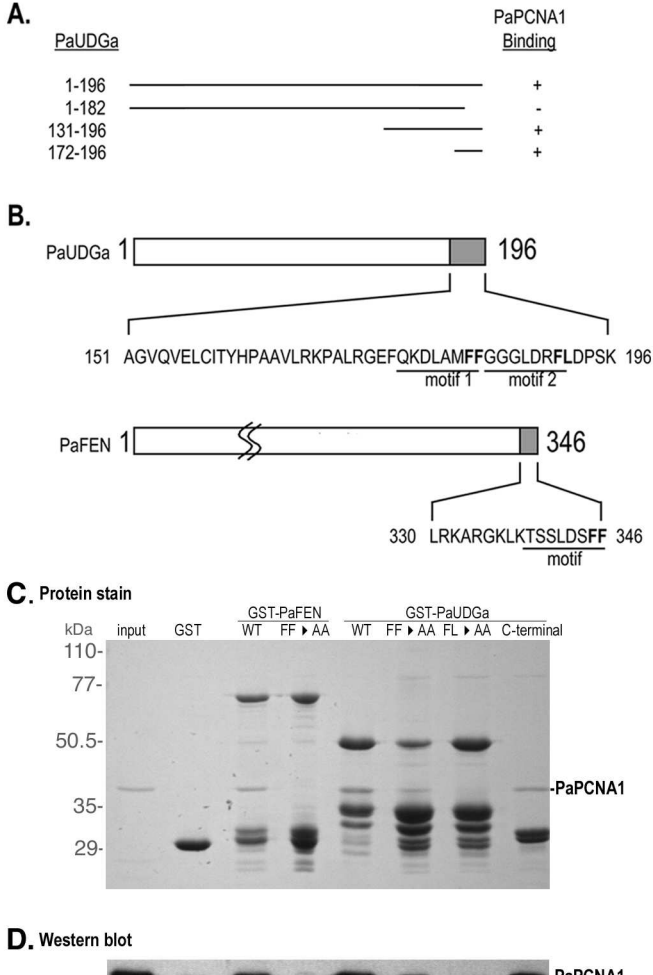


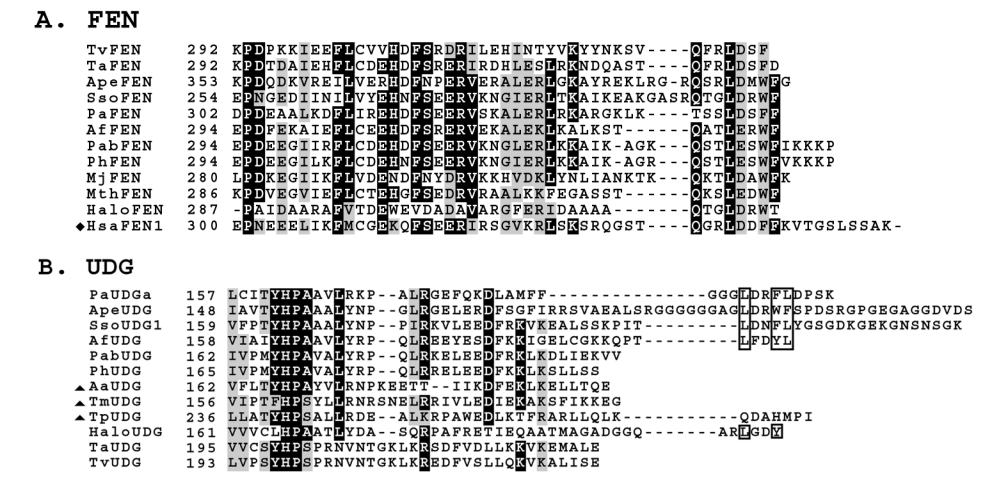
D. Protein stain



E. Protein stain




A.
PaUDGa
PaPCNA1 Binding
1-196 +
1-182 -
131-196 +
172-196 +
B.
PaUDGa 1 196
151 AGVQVELCITYHPAAVLRKPALRGEFQKDLAMFFGGGLDRFLDPSK 196
motif 1
motif 2
PaFEN 1 346
330 LRKARGKLKTSSLDSFF 346
motif
C. Protein stain
kDa
input
GST
GST-PaFEN
WT
FF ▸ AA
GST-PaUDGa
WT
FF ▸ AA
FL ▸ AA
C-terminal
110-
77-
50.5-
35-
29-
PaPCNA1
D. Western blot
PaPCNA1
lane 1 2 3 4 5 6 7 8


A. FEN
TvFEN 292 KPDPKKIEEFLCVVHDFSRDRILEHINTYVKYYNKSV----QFRLDSF
TaFEN 292 KPDTDAIEHFLCDEHDFSRERIRDHLESLRKNDQAST----QFRLDSFD
ApeFEN 353 KPDQDKVREIFVERHDFNPERVERALERLGKAYREKLRG-RQSRLDMWFG
SsoFEN 254 EPNGEDIINILTVYEHNFSEERVKNGIERLTKAIKEAKGASRQTGLDRWF
PaFEN 302 DPDEAALKDFIRIREHDFSEERVSKALERLRKARGKLK----TSSLDSFF
AfFEN 294 EPDFEKAIEFICEEHDFSRERVEKALEKLKALKST------QATLERWF
PabFEN 294 EPDEEGIIRFICDEHDFSEERVKNGLERLKKAIK-AGK---QSTLESWFIKKKP
PhFEN 294 EPDEEGILKFICDEHNFSEERVKNGIERLKKAIK-AGR---QSTLESWFVKKKP
MjFEN 280 LPDKEGIIKFIVDENDFNYDRVKKHVDKLYNLIANKTK---QKTLDAWFK
MthFEN 286 KPDVEGVIEFLCTEHGFSEDRVRAALKKFEGASST------QKSLEDWF
HaloFEN 287 -PAIDAARAFVTDEWEVDADAVARGFERIDAAAA------QTGLDRWT
◆HsaFEN1 300 EPNEEELIKFMCGEKQFSEERIRSGVKRLSKSRQGST----QGRLDDFFKVTGSLSSAK-
B. UDG
PaUDGa 157 LCITYHPAAVLRKP--ALRGEFQKDLAMFF----------GGGLDRFLDPSK
ApeUDG 148 IAVTYHPAAALYNP--GLRGELERDFSGFIRRSVAEALSRGGGGGGAGLDRWFSPDSRGPGEGAGGDVDS
SsoUDG1 159 VFPTYHPAAALYNP--PIRKVLEEDFRKVKEALSSKPIT---------LDNFLYGSGDKGEKGNSNSGK
AfUDG 158 VIAIYHPAAVLYRP--QLREEYESDFKKIGELCGKKQPT---------LFDYL
PabUDG 162 IVPMYHPAVALYRP--QLRKELEEDFRKLKDLIEKVV
PhUDG 165 IVPMYHPAVALYRP--QLRRELEEDFKKLKSLLSS
▲AaUDG 162 VFLTYHPAYVLRNPKEETT--IIKDFEKLKELLTQE
▲TmUDG 156 VIPTFHPSYLLRNRSNELRRIVLEDIEKAKSFIKKEG
▲TpUDG 236 LLATYHPSALLRDE--ALKRPAWEDLKTFRARLLQLK----------QDAHMPI
HaloUDG 161 VVVCLHPAATLYDA--SQRPAFRETIEQAATMAGADGQQ--------ARLGDV
TaUDG 195 VVCSYHPSPRNVNTGKLKRSDFVDLLKKVKEMALE
TvUDG 193 LVPSYHPSPRNVNTGKLKRSEDFVSLLQKVKALISE